

## SYSTEM AND PROCESS FOR RECOGNIZING AN AUDIO SEQUENCE

The invention concerns a process and a system for providing a user of a mobile telephony network with information or services relating to an audio sequence to which he is listening.

Persons who listen to a music segment, for example a piece broadcast over the radio or in a public place, sometimes want to obtain information concerning the piece to which they are listening. For example, they want to know the title or the performer of the piece that is broadcast in order that they might be able to obtain the piece of music.

The document WO 02/27600 (published on April 4, 2002) describes a system allowing a user to obtain information concerning an audio sequence to which he is listening. The audio sequence is transmitted to an interactive vocal server (IVR) by a capture device, for example a mobile terminal. The IVR server determines the characteristics ("fingerprint") of the audio sequence thus transmitted. It compares these characteristics with characteristics stored in a database and associated with predetermined audio sequences. On the basis of this comparison, the IVR server sends back to the user information relating to the identified audio sequence. This system also allows the user to obtain additional services, such as, for example, buying the identified piece of music.

With such a system, the user must be able to establish a connection with the IVR server from his mobile terminal, for example by dialing a special number. Moreover, he must transmit the audio sequence to which he is listening to the server in real time. The user must also be capable of interacting with the IVR server to obtain additional services.

Moreover, the IVR server is a complex server, since it must be able to exchange voice data with the user and process it.

One goal of the invention is to propose a system allowing assisted access to audio sequence recognition services.

To accomplish this, the invention proposes a mobile terminal characterized in that it comprises means of management containing an application that is able to execute the following steps automatically:

- command the establishment of a connection between the mobile terminal and a remote audio sequence recognition server;
- after the connection is established, command the transmission of an audio sequence to the server, in order for the server to identify said audio sequence.

The mobile terminal is able to make the connection to the remote audio sequence recognition server automatically, and to transmit the audio sequence to be identified, so that the user need not perform any specific manipulations. In particular, he need not dial any special connection number. The audio sequence recognition service is easy for the user to access.

Moreover, the user interacts with his terminal and not with the remote server. Interaction with the remote server is managed by means of management within the terminal. Access to the audio sequence recognition service is presented to the user in an ergonomic manner.

In one embodiment of the invention, the terminal comprises means of memory that are able to record an audio sequence, and the application is able to command the means of memory to record the audio sequence before commanding the establishment of a connection between the mobile terminal and the remote audio sequence recognition server. Thus, in case the application encounters difficulties establishing a connection with the remote audio sequence recognition server, it can automatically renew the connection later.

Moreover, the terminal can convert the recorded audio sequence in order to transmit it to the remote audio sequence recognition server in an appropriate form so that it can be identified by the server, for example in the form of data packets.

In this embodiment, the audio sequence recognition server does not require means to exchange and process voice data.

In another embodiment of the invention, the application is able to transmit the audio sequence in real time by audio streaming (continuous transmission of the audio sequence as it is generated).

In another embodiment of the invention, the application is able to command the means of management to determine a signature of the audio sequence, and to transmit this signature in

real time by audio streaming (continuous transmission of the signature as it is generated).

In another embodiment of the invention, the terminal comprises means of memory that are able to record an audio sequence, and the application is able to command the means of memory to record the audio sequence. The application is then able to command the means of management to determine a signature of the audio sequence in order to transmit this signature to the audio sequence recognition server.

In this embodiment, the audio sequence is transmitted to the audio sequence recognition server directly in an appropriate form (signature) for the server to perform its identification. The audio sequence recognition server does not require means to convert the audio sequence into a signature.

The invention also relates to an audio sequence recognition process characterized in that it comprises launching an application contained in means of management of a mobile terminal, said application automatically performing the following steps:

- commanding establishment of a connection between the mobile terminal and a remote audio sequence recognition server;
- after the connection is established, commanding the transmission of an audio sequence to the server, in order for the server to identify said audio sequence.

Other characteristics and advantages of the invention will still follow from the description which follows. This description illustrates a possible embodiment of the invention. It should be read in conjunction with the attached figures, among which:

- Figure 1 is a schematic representation of an embodiment of an audio sequence recognition system according to the invention;
- Figure 2 is a schematic representation of the different steps of the audio sequence recognition process according to one embodiment of the invention.

In Figure 1, the audio sequence recognition system comprises a mobile terminal 100, a remote audio sequence recognition server 200, and a services server 300.

Mobile terminal 100 comprises means 110 for capturing an audio sequence, in the form of a microphone, means of memory 130 for recording the captured audio sequence, means of

management 120 in the form of a microprocessor, and means of transmission / reception 140, in the form of an antenna. Means of management 120 contain an application allowing automatic initiation of the various steps of audio sequence recognition.

Means of management 120 command the screen 150 of mobile terminal 100 to display a link corresponding to the audio sequence recognition service. The link is presented in the form of a specific icon that the user can select to launch the application. When the link is activated by the user, the means of management 130 start the application.

Audio sequence recognition server 200 is connected to a database 210 containing a set of audio sequence signatures as well as identification data associated with each of these audio sequences.

The service server 300 is able to perform services relating to the identified audio sequences, and to record the data necessary for billing these services. These services consist, for example, of obtaining information relating to the identified audio sequence, purchasing mobile content associated with the identified audio sequence, or purchasing a product relating to the identified audio sequence. To accomplish this, the services server 300 is connected to a group of specialized servers 310, 320, 330. For example, specialized server 310 is connected to a database of mobile content and is able to provide mobile content on the basis of identification information that is transmitted to it.

The audio recognition process will be described in relation to Figures 1 and 2.

A user hears an audio sequence S which interests him and wants to obtain information or services relating to this audio sequence. In a first step 10, the user launches the audio sequence recognition application by selecting the corresponding icon on his mobile terminal 100. Selection of the icon has the effect of starting the audio sequence recognition application, which automatically performs the following steps.

In a step 20, the application commands means of memory 130 of the mobile terminal 100 to record the audio sequence S.

In a step 30, the application continues recording for a predetermined recording interval, for

example 15 seconds. This predetermined recording interval depends on the recognition performance characteristics that are sought.

In a step 40, after said predetermined recording interval has passed, the application commands the establishment of an Internet connection between the mobile terminal 100 and the remote audio sequence recognition server 200.

In a step 50, when the connection with the server has been established, the application commands transmission of the audio sequence that was recorded in means of memory 130 to the remote audio sequence recognition server 200. To accomplish this, the audio sequence can be compressed beforehand and transmitted in the form of data packets. Alternatively, the audio sequence can be transmitted by audio streaming.

The audio sequence recognition server 200 determines a signature from the audio sequence thus transmitted, and compares this signature with the signatures contained in database 210. If the result of this comparison is positive, that is if server 200 has found a corresponding signature among the signatures contained in the database, server 200 sends a signal to mobile terminal 100 indicating that it has identified the audio sequence containing a reference associated with the audio sequence thus identified.

In a step 60, the mobile terminal 100 waits to receive the identification signal from the audio sequence recognition server 200.

In a step 70, when the mobile terminal has received the identification signal, the application commands the recording of the reference associated with the audio sequence in the means of memory 130 and displays, on screen 150 of mobile terminal 100, a menu comprising a series of proposed services relating to the identified audio sequence which can be selected by the user.

The displayed menu comprises the following choices:

- 1) obtain information relating to the identified audio sequence;
- 2) purchase mobile content associated with the identified audio sequence;
- 3) purchase a product relating to the identified audio sequence.

In a step 80, the user selects one of the proposed services that are displayed.

If the user selects proposed service 1), this signifies that he wants to obtain identification information relating to the audio sequence. For example, if the audio sequence is a song, he could obtain the title, the performer, the name of the album containing the song, the price of the album, and any other useful information.

If the user selects proposed service 2), this signifies that he wants to purchase mobile content associated with the identified audio sequence. For example, if the audio sequence is a song, he could obtain a ringtone and wallpaper corresponding to the song.

If the user selects proposed service 3), this signifies that he wants to purchase a product relating to the identified audio sequence, for example a disk.

In step 90, the application commands the sending of a message containing a request corresponding to the selected proposed service to the service server 300 in order for the corresponding service to be performed. The message also contains the reference associated with the identified audio sequence. The message can be sent in the form of an SMS message or any other appropriate form.

Suppose, for example, that the user selected proposed service 2) corresponding to obtaining mobile content. The service server 300 receives the request, and transmits the reference to the mobile content server 310. On the basis of the reference, the content server 310 is able to retrieve the corresponding mobile content from the database 350.

The service server 300 transmits the mobile content to the terminal 100. The service server 300 is also able to perform different authentication and recording operations to allow billing of the mobile content transmission to the user.

It is advantageous for the means of memory 130 of the mobile terminal 100 to contain the references of the last ten audio sequences identified by the audio sequence recognition server 200 at the request of the user. Thus, every time the user launches the application, he has direct access to the services relating to the audio sequences already identified.

One embodiment of the audio sequence recognition process has been described in which the application performs steps 20 and 30, which consist of the means of memory 130 recording the audio sequence, before a command is issued to establish a connection between the mobile terminal 100 and the remote audio sequence recognition server 200.

In another embodiment, the audio sequence to be identified is not recorded in means of memory before being transmitted to the audio sequence recognition server 200. The sequence is transmitted in real time to the audio sequence recognition server by audio streaming (continuous transmission of the audio sequence as it is captured by the microphone).

In yet another embodiment, the application ~~the audio sequence to be identified is not recorded in means of memory before being transmitted to the audio sequence recognition server 200.~~ The application commands the means of management 120 to determine a signature of the audio sequence and transmit this signature to the audio sequence recognition server 200, as this signature is generated.

In yet another embodiment, the application performs steps 20 and 30, which consist of recording the audio sequence in the means of memory 130, and then it commands the means of management 120 to determine a signature of the recorded audio sequence and transmit this signature to the audio sequence recognition server 200.

In these two last embodiments, the audio sequence is transmitted to the audio sequence recognition server 200 in an appropriate form for the server to compare it directly with the signatures contained in database 210. In this embodiment, part of the recognition process is performed by the terminal, so that the audio sequence recognition server is relieved of this part of the process. Moreover, once the audio sequence has been converted into a signature, the bandwidth that is necessary to transmit the converted audio sequence is smaller than that which is necessary to transmit the unconverted audio sequence directly.

Moreover, the invention can allow the user to provide the audio sequence recognition server 200 with data relating to the audio sequence in which he is interested, to facilitate its identification. For example, the user can use the recognition system when what interests him about the audio sequence is not identifying it as such, but the different services associated with this audio sequence. It is advantageous if the data supplied by the user can be used by

the audio sequence recognition server 200 to complete the database 210 or confirm / modify the information that it already contains.